SL

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/938,104	08/23/2001	Adam Strauss	42390P12534	1037

8791 7590 09/02/2004

BLAKELY SOKOLOFF TAYLOR & ZAFMAN
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1030

EXAMINER	
WOZNIAK, JAMES S	

ART UNIT	PAPER NUMBER
2655	9

DATE MAILED: 09/02/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

Office Action Summary	Application No.	Applicant(s)	
	09/938,104	STRAUSS ET AL.	
	Examiner	Art Unit	
	James S. Wozniak	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on 8/23/2001.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-53 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-53 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 31 December 2001 is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____.

Detailed Action

Drawings

1. New corrected drawings are required in this application because the current drawings are informal.

The requirement for corrected drawings will not be held in abeyance.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. **Claims 1, 3, 7-9, 11, 13, 17-19, 21, 22, 24, 28-30, 32, 33, 35, 39-41, 43, 44, 46, and 50-52** are rejected under 35 U.S.C. 103(a) as being unpatentable over Gupta et al *(U.S. Patent: 5,459,814)*.

With respect to **Claims 1 and 11**, Gupta discloses:

An integrated voice activation detector and method for detecting whether voice is present, the integrated voice activation detector comprising:

A semiconductor integrated circuit including,

At least one signal processing unit to perform voice detection *(multiple DSPs, Fig. 1, and Col. 2, Lines 45-53)*; and

A processor readable storage means to store signal-processing instructions for execution *(DSP memory, Fig. 1, and Col. 2, Lines 50-53)* by the at least one signal processing unit to:

Detect whether noise is present to determine whether a noise flag should be set *(detecting background noise based upon a signal energy level, Col. 3, Lines 53-58, and a lower level threshold comparison for detecting noise, Col. 5, Lines 20-23, and Fig. 4, Element 33)*;

Detect a predetermined number of zero crossings to determine whether a zero crossing flag should be set *(zero crossings, Col. 3, Line 66- Col. 4, Line 5, and zero crossing threshold comparison, Col. 5, Lines 24-26 and Fig. 4, Element 35)*;

Detect whether a threshold amount of energy is present to determine whether an energy flag should be set *(signal energy level, Col. 4, Lines 17-19, to an upper level threshold for speech detection, Col. 5, Lines 18-21 and Fig. 4, Element 31)*;

Detect whether instantaneous energy is present to determine whether a instantaneous energy flag should be set *(detecting rapid changes in a signal energy level through a slope measurement, Col. 3, Lines 59-65, and a slope measurement threshold comparison, Col. 5, Lines 28-31, and Fig. 4, Element 37)*; and

Utilize a combination of the noise, zero crossing, energy, and instantaneous energy flags to determine whether voice is present *(Fig. 4)*.

Although Gupta teaches setting a VAD flag based on a noise, zero crossing, energy, or slope threshold comparison result as seen in Fig. 4, Gupta does not specifically disclose setting intermediate flags corresponding to the aforementioned threshold comparisons, however, the

examiner takes official notice that it is well known in the art to set flags in a DSP to indicate the result of a processor calculation (as is evidenced by the VAD flag taught by Gupta) so that further appropriate signal processing can be implemented. Therefore, it would have been obvious to one of ordinary skill in the art, at the time of invention, that the DSP taught by Gupta would set flag bits based on the noise, zero crossing, energy, and slope threshold comparison test results for further analysis by the processor to determine the presence of speech, indicated by setting a VAD flag.

With respect to **Claims 3, 13, 24, 35, and 46**, Gupta discloses:

Interim voice activity decision flag being set to indicate voice has been detected by determining if the instantaneous energy flag is set or the energy flag is set and the noise flag is not set and the zero crossing flag is not set *(VAD flag set to 1, Fig. 4, Element 38)*.

With respect to **Claims 7, 17, 28, 39, and 50**, Gupta discloses:

Detecting a predetermined number of zero crossings to determine whether a zero crossing flag should be set includes determining whether a root mean square crossing value is greater than a threshold value *(zero crossing threshold comparison, Fig. 4, Element 35, and Col. 5, Lines 24-26)*.

Although Gupta does not specifically disclose that an rms value is compared to a threshold for a zero crossing threshold comparison, the examiner takes official notice that a rms measurement is a means well known in the art for representing signal energy. Therefore, it would have been obvious to one of ordinary skill in the art, at the time of invention, to utilize the well-known rms energy measurement as a means of expressing a signal level with respect to a zero crossing point in order to determine a zero crossing result for threshold comparison.

With respect to **Claims 8, 18, 29, 40, and 51**, Gupta recites:

Detecting whether noise is present to determine whether a noise flag should be set includes determining whether energy in a current frame multiplied by a threshold is greater than delayed frame energy *(comparing a signal energy level to a lower level threshold to detect the presence of noise, Col. 5, Lines 20-23, and threshold adjustment based upon past noise energy levels, Col. 6, Lines 18-39)*.

With respect to **Claims 9, 19, 30, 41, and 52**, Gupta teaches the VAD device and method utilizing noise, zero crossing, energy, and slope threshold comparisons in determining the presence of speech, as applied to Claims 1 and 11. Gupta does not teach the use of a autocorrelation logarithm in determining if speech is present through threshold comparison, however, the examiner takes official notice that it is well known in the art that an autocorrelation involves an energy measurement and is often used in the art as a means of expressing a signal energy level. Therefore, it would have been obvious to one of ordinary skill in the art, at the time of invention, to use an logarithm of an autocorrelation as a well-known means of expressing an energy level for threshold comparison in speech detection since speech data tends to correlate over a wider range than noise, thus the greater the value of an autocorrelation, the higher the likelihood of speech presence.

Claims 21, 22, 32, 33, 43, and 44 contain subject matter similar to Claim 1, and thus, are rejected for the same reasons.

4. **Claims 2, 12, 23, 34, and 45** are rejected under 35 U.S.C. 103(a) as being unpatentable over Gupta et al in view of Galand et al *(U.S. Patent: 4,782,523)*.

With respect to **Claims 2, 12, 23, 34, and 45**, Gupta teaches the VAD device and method utilizing noise, zero crossing, energy, and slope threshold comparisons in determining the presence of speech, as applied to Claims 1 and 11. Gupta does not disclose the use of an FFT to determine whether an FFT flag should be set, however the use of an FFT is well-known in the telephony art for DTMF signal detection, as is evidenced by Garland:

The signal processing instructions further for execution by the at least one signal processing unit to, perform fast Fourier transformation (FFT) processing *(FFT approach for tone detection, Col. 2, Lines 24-28)*.

Gupta and Galand are analogous art because they are from a similar field of endeavor in signal detection for telephonic communications. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to combine the use of an FFT for tone signal detection as taught by Galand with the VAD device and method utilizing noise, zero crossing, energy, and slope threshold comparisons in determining the presence of speech as taught by Gupta to provide for signal detection of a tone signal by a VAD (since this tone would not be correctly classified as noise or speech) to ensure that a tone signal would not be confused with a voice signal in a telephonic application such as a speech driven menu system that is capable of also accepting a DTMF input. Also, it would also have been obvious to utilize a flag to indicate the detection of such a tone signal, as per the reasons given for flag usage with respect to Claim 1. Therefore, it would have been obvious to combine Galand with Gupta for the benefit of

obtaining tone detecting means in a VAD, to obtain the invention as specified in Claims 2, 12, 23, 34, and 45.

5. **Claims 4-6, 14-16, 25-27, 36-38, and 47-49** are rejected under 35 U.S.C. 103(a) as being unpatentable over Gupta et al in view of Kapanen *(U.S. Patent: 5,835,889)*.

With respect to **Claims 4, 14, 25, 36, and 47**, Gupta teaches the VAD device and method utilizing noise, zero crossing, energy, and slope threshold comparisons in determining the presence of speech, indicated by a VAD flag, as applied to Claims 3 and 13. Although Gupta does disclose the consideration of a past flag value and hangover processing for threshold adjustment as shown in Fig. 2, Gupta does not teach that a hangover calculation is applied to determining whether to set or clear a VAD flag, however Kapanen recites:

Perform HangOver and Speech Kick in processing after the interim voice activity decision has been made to determine whether a voice activity flag should be set or cleared *(resetting a speech detection flag only after a hangover period has elapsed, Col. 5, Lines 20-30)*.

Gupta and Kapanen are analogous art because they are from a similar field of endeavor in voice activity detection. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to combine the use of a hangover period in VAD as taught by Kapanen with the VAD device and method utilizing noise, zero crossing, energy, and slope threshold comparisons in determining the presence of speech, indicated by a VAD flag as taught by Gupta to improving VAD accuracy by ensuring that a signal contains only noise and that speech from a user has completely ceased before a VAD determines that speech is not present and clears the corresponding flag. Also, since the same principles can be applied to a speech signal start after a

noise period, it would have been obvious to one of ordinary skill in the art, at the time of invention, to utilize a noise hangover period to ensure that valid speech data has begun instead of a random noise spike before setting a VAD flag to 1. Therefore, it would have been obvious to combine Kapanen with Gupta for the benefit of obtaining a VAD capable of further ensuring a valid noise or speech signal through the use of a hangover period, to obtain the invention as specified in Claims 4, 14, 25, 36, and 47.

With respect to **Claims 5, 15, 26, 37, and 48**, Gupta further discloses:

If the voice activity flag is set, send a speech payload to be packetized and update the voice activity detection flag for external interaction with other functions of the semiconductor integrated circuit *(CELP coder as a DSP capable of sending and receiving speech data, Fig. 1. Also, the VAD flag would be updated upon reception of speech data as per the feedback loop noted above with respect to Claim 4).*

With respect to **Claims 6, 16, 27, 38, and 49**, Gupta further discloses:

If the voice activity flag is not set, disable an automatic level control and cause a silence insertion description payload to be prepared (*CELP coder as a DSP capable of sending and receiving speech data that would include unvoiced speech, Fig. 1 and unvoiced speech, Col. 3, Lines 46-47).*

Although Gupta does teach a best gain calculation for speech data, Col. 3, Lines 2-8, Gupta does not specifically suggest that the gain is calculated using an automatic gain control, however, the examiner takes official notice that it is well known in the art to utilize a means for automatic gain control in CELP coding in order to maintain an acceptable perceptible speech level upon reception. Therefore, it would have been obvious to one of ordinary skill in the art, at

the time of invention, to utilize automatic gain control in the CELP gain calculation taught by Gupta in order to maintain a perceptible speech signal level upon reception. Also, since the speech data in this case contains only silence and no speech information of value, no signal amplification or attenuation would be necessary, thus, it would have been obvious to disable the automatic gain control.

6. **Claims 10, 20, 31, 42, and 53** are rejected under 35 U.S.C. 103(a) as being unpatentable over Gupta et al in view of Atal et al ("A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Applications to Speech Recognition," 1976).

With respect to **Claims 10, 20, 31, 42, and 53**, Gupta teaches the VAD device and method utilizing noise, zero crossing, energy, and slope threshold comparisons in determining the presence of speech, as applied to Claims 1 and 11. Gupta does not teach the use of an autocorrelation difference at a delayed sample for threshold comparison to detect instantaneous speech energy, however, such a comparison method is well known in the art as is evidenced by Atal:

Detecting whether instantaneous energy is present to determine whether an instantaneous energy flag should be set includes determining whether a difference between a current frames energy at an autocorrelation of a tenth delayed sample and a prior frames energy at an autocorrelation of a tenth delayed sample is greater than a previous frames autocorrelation multiplied by a threshold *(detecting speech through the use of an autocorrelation coefficient at a unit sample delay, Page 202, Section II)*.

Gupta and Atal are analogous art because they are from a similar field of endeavor in voice activity detection. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to combine the use of an autocorrelation at a unit delay for speech detection as taught by Atal with the VAD device and method utilizing noise, zero crossing, energy, and slope threshold comparisons in determining the presence of speech as taught by Gupta to provide a well-known means of detecting instantaneous energy changes, representative of speech, through the use of an autocorrelation at a delay since speech is correlated over a wider range than noise, a high autocorrelation at a delay would be indicative of speech presence.

Also, Atal does not specifically teach the use of an autocorrelation of a tenth delayed sample , however, it would have been obvious matter of design choice to utilize the autocorrelation of a tenth delayed sample for speech detection, since the applicant has not disclosed that specifically using an autocorrelation of a tenth delayed sample solves any stated problem or is for any particular purpose. An autocorrelation of a tenth delayed sample would provide enough delay to sufficiently indicate the presence of speech, which correlates over a wider range than noise, and thus would be an obvious choice for the unit sample delay taught by Atal.

Therefore, it would have been obvious to combine Atal with Gupta for the benefit of obtaining a means of detecting speech through a threshold comparison of an autocorrelation sample at a unit delay that is capable of sufficiently indicating the presence of speech, to obtain the invention as specified in Claims 10, 20, 31, 42, and 53.

Conclusion

7. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

- Swaminathan et al *(U.S. Patent: 5,596,676)*- discloses a speech coder that determines the presence of speech based upon zero-crossing and pitch flags.
- Chiba et al (U.S. Patent: 5,727,121)- teaches a speech detection method that implements a short-term energy threshold comparison.
- Benyassine et al *(U.S. Patent: 5,774,849)*- teaches a voice activity detector that utilizes energy, zero-crossing, and autocorrelation data in making a voicing decision.
- Sonnic *(U.S. Patent: 6,154,721)*- discloses a VAD that compares zero-crossings and signal energy to a threshold for speech detection and counts the number of consecutive speech or noise frames.

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James S. Wozniak whose telephone number is (703) 305-8669 and email is James.Wozniak@uspto.gov. The examiner can normally be reached on Mondays-Fridays, 8:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Ivars Smits can be reached at (703) 306-3011. The fax/phone number for the Technology Center 2600 where this application is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the technology center receptionist whose telephone number is (703) 306-0377.

James S. Wozniak
8/23/2004

W. R. YOUNG
PRIMARY EXAMINER